

## Research Article

# Automatic Sex Determination of Skulls Based on a Statistical Shape Model

**Li Luo,<sup>1</sup> Mengyang Wang,<sup>1</sup> Yun Tian,<sup>1</sup> Fuqing Duan,<sup>1</sup> Zhongke Wu,<sup>1</sup> Mingquan Zhou,<sup>1</sup> and Yves Rozenholc<sup>2</sup>**

<sup>1</sup> College of Information Science and Technology, Beijing Normal University, Beijing 100875, China

<sup>2</sup> Laboratoire MAP5, UMR CNRS 8145, University Paris Descartes, 45 rue des Saints-Pères, 75006 Paris, France

Correspondence should be addressed to Fuqing Duan; [fqduan@bnu.edu.cn](mailto:fqduan@bnu.edu.cn)

Received 11 July 2013; Accepted 4 October 2013

Academic Editor: Yunmei Chen

Copyright © 2013 Li Luo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sex determination from skeletons is an important research subject in forensic medicine. Previous skeletal sex assessments are through subjective visual analysis by anthropologists or metric analysis of sexually dimorphic features. In this work, we present an automatic sex determination method for 3D digital skulls, in which a statistical shape model for skulls is constructed, which projects the high-dimensional skull data into a low-dimensional shape space, and Fisher discriminant analysis is used to classify skulls in the shape space. This method combines the advantages of metrical and morphological methods. It is easy to use without professional qualification and tedious manual measurement. With a group of Chinese skulls including 127 males and 81 females, we choose 92 males and 58 females to establish the discriminant model and validate the model with the other skulls. The correct rate is 95.7% and 91.4% for females and males, respectively. Leave-one-out test also shows that the method has a high accuracy.

## 1. Introduction

Sex identification from skeletons is a vital work for a forensic anthropological analysis. Previous studies [1–4] indicate that pelvis is the most reliable indicator of sex assessment, and skull is the second one. However, not all forensic cases provide a complete skeleton due to breakage or postmortem destruction, while the skull can be well preserved in most cases since it is composed of hard tissue. So the skull is the most commonly used skeleton part in forensic anthropological analysis.

Traditional skeletal sex assessments principally rely on visual assessments of sexually dimorphic traits. Rogers [5] achieves an accuracy of 89.1% for a historic skeletal collection by using visual morphological traits. By scoring the visual assessments of five cranial traits (glabella, mental, orbit, nuchal, and mastoid), Walker [6] achieves 90% accuracy for modern American skulls via a quadratic discriminant analysis model incorporating scores. According to Daubert [7] and Mohan [8] criteria, Williams and Rogers [9] assess 21 skull characteristics of 50 white European Americans (25 males and 25 females). For a characteristic, if the intraobserver error

is no more than 10% and the accuracy is above 80% when it is used separately to identify the sex, the characteristic is defined as a high quality characteristic. They get six high quality characteristics like eyebrow, orbit, and so forth, and the accuracy reaches 94% by utilizing visual assessments of all these six characteristics. Visual assessments depend heavily on physical anthropologists' understanding of population differences in sexual dimorphism. The visual assessment results reported in the paper [10] show that Krogman and Iscan achieve 92% accuracy using the Todd collection, while Stewart only obtains 77% accuracy using some American black skulls from the Terry collection. Ramsthaler et al. [11] use kappa statistics to quantify the disagreement in sex classification performed by two different observers after visual assessment, and the agreement only reaches 90.8%. Moreover, visual assessment of the morphological traits is likely to be inaccurate when performed by an inexperienced observer due to its great subjectivity. With the progress of the digital imaging technology in medicine, discriminant analysis for skeletal measurements is increasingly used for sex estimation.

Since radiograph can provide architectural and morphological details of the skull, some researchers [12–15] use skull radiograph for sex identification. Hsiao et al. [12] consider the lateral radiographs of 50 male and 50 female adult skulls from Taiwan. They use 18 cephalometric variables, which are derived from some cephalometric points on the lateral radiographs, to build the discriminant function, and claim a coincidence rate of 100%. But they do not mention the generalization ability of the discriminant function, that is, the classification ability for new samples. Veyre-Goulet et al. [15] validate the method of Hsiao by using the lateral radiographic data of 114 Europeans. They find that the result using 8 cephalometric variables is the same as the one using 18 variables, and the coincidence rate is 95.6%. They think that different cephalometric variables can be used for different races. Inoue et al. [14] compute the gradients and distances of 39 measuring points in the lateral contour of the skull and establish the discriminant function through these variables by using 50 female and 50 male skulls from Japan as specimens. 21 other specimens are tested, and the mean correct rate is 86%. They conclude that gradient has higher ability in reflecting gender differences than distances.

Other researchers use some variables measured from 3D skulls to establish discriminant functions. Spradley and Jantz [16] construct multivariate discriminant models using the skull data from the Forensic Anthropology Data Bank. They measure different variables for different races, and the highest identification rate is no more than 90%. Jantz and Ousley [17] issue the computer-aided software called Fordisc for skeleton analysis. This software measures skeleton characteristics by the human-computer interactive way and utilize the data from American Forensic Data Bank to establish the discriminant function. Guyomarc'h and Bruzek [18] compare the sex identification effects of the software for different races such as Thailand and France, and the correct recognition rate is between 52.2% and 77.8%. Ramsthaler et al. [19] test this software by using 98 Caucasian German skulls, and the correct rate is 86%. These studies show that the craniofacial morphology has great diversities among different races. For Chinese, Li [20] manually extracts the mid-sagittal frontal arc on dried skulls from northeast China, and the Fourier coefficients of the arc are used for multivariate stepwise discriminant analysis. The recognition rate for 31 test skulls is 84.21% and 83.33% for males and females, respectively. Shui [21] measures 14 metric variables and utilizes 94 skulls from north China to establish stepwise fisher discriminant functions. The recognition rate for 39 test skulls reaches 87%.

Discriminant analysis for skeletal measurements requires a high measurement precision. However, the accurate measurement of the skull is quite difficult. Williams and Rogers [9] show that, for most measure variables, usually the measurement error among different observers is above 10%. Moreover, the skull size changes with varying ages. In this work, we propose an automatic sex determination method for 3D digital skulls. A statistical shape model is established to describe statistical features of the skull morphology, and Fisher discriminant analysis of the shape parameters is used to classify skulls. The advantages of this method is as follows: Firstly, it needs no professional qualification and tedious

manual measurement; Secondly, it is less influenced by the variation of the skull size; Finally, it can get a high recognition rate.

## 2. Materials and Methods

**2.1. Materials.** This study has been approved by the Institutional Review Board (IRB) of Image Center for Brain Research, National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University. It is carried out on a database of 208 whole skull CT scans on voluntary persons that mostly come from Han ethnic group in North of China, age 19–75 years for females and 21–67 years for males. There are 81 females and 127 males. The CT images were obtained by a clinical multislice CT scanner system (Siemens Sensation 16) in the Xianyang hospital located in western China. The images of each subject are restored in DICOM format with a size of approximately  $512 \times 512 \times 250$ . Each 3D skull surface is extracted from the CT images and represented as a triangle mesh including about 150,000 vertices. All the skulls are substantially complete; that is, each skull contains all the bones from calvarias to jaw and has the full mouth of teeth.

To eliminate the inconsistency in position, pose, and scale caused by data acquirement, all the samples are transformed into a uniform coordinate system. The uniform coordinate system is determined by four skull landmarks, left porion, right porion, left (or right) orbitale, and glabella (denoted as  $L_p, R_p, L_o, G$ ). From three points,  $L_p, R_p, L_o$ , the Frankfurt plane is determined [22]. The coordinate origin (denotes as  $O$ ) is the intersection point of the line  $L_p R_p$  and the plane that contains point  $G$  and orthogonally intersects with line  $L_p R_p$ . We take the line  $OR_p$  as  $x$ -axis. The  $z$ -axis is the line through the point  $O$  and with the direction being the normal of the Frankfurt plane. Then  $y$ -axis is obtained by the cross product of  $z$ - and  $x$ -axis. Once the uniform coordinate system is defined, all the prototypic skulls are transformed into it. Finally, the scale of all the samples is standardized by setting the distance between  $L_p$  and  $R_p$  to unit; that is, each vertex  $(x, y, z)$  of the skull is scaled by  $(x/|L_p - R_p|, y/|L_p - R_p|, z/|L_p - R_p|)$ . One skull in the uniform coordinate system is shown in Figure 1.

**2.2. Statistical Shape Model Construction.** Statistical shape model is a widely used technique in medical image analysis. It can efficiently describe the shape variance and ensure that only statistically likely shapes are represented. Principal Component Analysis (PCA) [23] is a powerful tool to build statistical shape models, and it finds the major and minor modes of shape variation across the training data. In order to build the statistical model, a dense point correspondence has to be established across the training set, that is, building a point-to-point correspondence for all training samples according to human anatomy characteristics. There are many nonrigid registration methods [24–26] for dense mesh or point cloud objects, and here we adopt the method presented by Hu et al. [25]. As in [25], the back part of the reference skull is cut away (as shown in Figure 2) considering that there

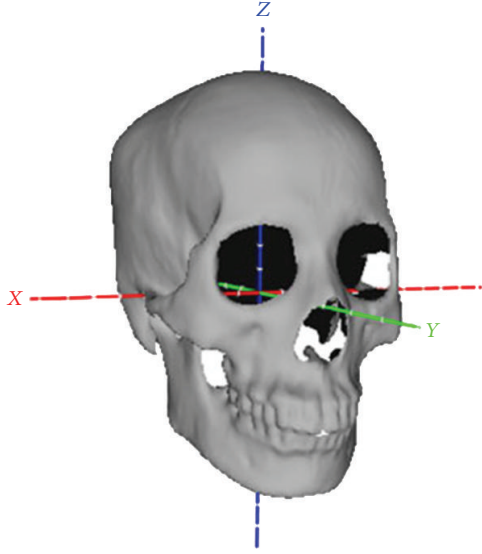


FIGURE 1: One skull in the uniform coordinate system.



FIGURE 2: The back part of the reference skull is cut away.

are too many vertices in the whole skull, and the sexually dimorphic traits are mainly on the front part of skulls. So all of the aligned training skulls only keep the front part corresponding to the reference.

By concatenating coordinates of all the vertices, a skull can be represented as a high dimension vector. Thus we construct a dataset of skulls  $\{\mathbf{S}_i = (x_{i1}^S, y_{i1}^S, z_{i1}^S, \dots, x_{im}^S, y_{im}^S, z_{im}^S)^T \mid i = 1, 2, \dots, N\}$ , where each coordinate index labels corresponding points across the training set.

From the skull dataset, the mean skull data  $\bar{\mathbf{S}}$  and covariance matrix  $\Sigma$  of the mean normalized skulls are calculated. PCA essentially transforms the mean normalized shape data into a subspace spanned by the orthogonal unit eigenvectors  $\mathbf{U}_k$ ,  $k = 1, 2, \dots, N - 1$ , of the covariance matrix in descending order according to their associated eigenvalues

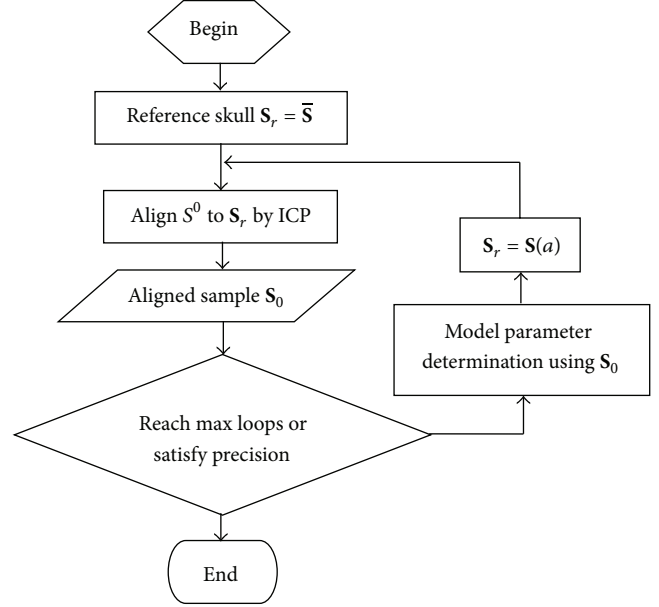


FIGURE 3: Model matching procedure.

$\lambda_k$ , which represent variation modes of the data. Then the skull statistical shape model is constructed as in the following parameterized model:

$$\mathbf{S}(\mathbf{a}) = \bar{\mathbf{S}} + \sum_{k=1}^M a_k \mathbf{U}_k, \quad (1)$$

where the number  $M$  is the mode number, usually determined by a variance contribution rate calculated from the cumulative eigenvalues, and the combination coefficient  $\mathbf{a} = (a_1, a_2, \dots, a_M)^T$  is the model parameter. Apparently this statistical shape model assumes that the shape vectors  $\mathbf{S}$  obey a normal distribution with mean  $\bar{\mathbf{S}}$  and covariance matrix  $\Sigma$ , so the parameter  $\mathbf{a}$  for a plausible skull meets a normal distribution with zero mean and covariance matrix  $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ .

**2.3. Statistical Shape Model Matching.** Model matching is to determine the model parameter for a given skull. If the skull data  $\mathbf{S}_0$  is aligned, the model parameters can be determined by PCA transform. According to the statistical models (1), let  $\mathbf{P}_s = [\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_M]$  denote the PCA transform matrix for skulls, and the model parameter can be determined as follows:

$$\mathbf{a} = \mathbf{P}_s^T (\mathbf{S}_0 - \bar{\mathbf{S}}). \quad (2)$$

So in fact the model matching is a procedure of the skull registration. Given an unknown prototypic skull data  $\mathbf{S}^0$ , it is firstly transformed into the uniform coordinate system as Section 2.1 describes. Figure 3 shows a statistical shape model based registration algorithm. As shown in Figure 3, a dynamic reference, denoted as  $\mathbf{S}_r$ , is updated by the statistical shape model in each loop, whose model parameter is determined by the PCA transform of the corresponding aligned sample

$S_0$  of last iteration. The initial reference is selected as the mean shape, and an ICP algorithm [27] is used to align a target to the reference. Apparently, the dynamic reference will be closer and closer to the target skull along with the iterating, so the iteration will converge. When the aligned sample does not change, the iteration stops.

**2.4. Fisher Discriminant Analysis in the Shape Parameter Space.** Fisher discriminant analysis is increasingly used for skull sex determination. It projects samples from a high-dimensional feature space into one axis called Fisher vector, in which optimal linear classification can be achieved. Different from previous methods, we perform Fisher discriminant analysis not for skull measurements, but for shape parameters of skulls.

According to the statistical shape model (1), a skull is represented as a feature vector  $\mathbf{a} = (a_1, a_2, \dots, a_M)^T$ . Let  $\mathbf{m}_i$ ,  $n^i$ ,  $i = 1, 2$ , denote the mean and the number of the training samples of the  $i$ th class, and let  $\mathbf{X}_j^i$  be the feature vector of the  $j$ th sample of the  $i$ th class. Then within-class scatter matrix and between-class scatter matrix can be defined as follows:

$$\mathbf{S}_w = \sum_{i=1}^2 \left( \sum_{j=1}^{n^i} (\mathbf{X}_j^i - \mathbf{m}_i) (\mathbf{X}_j^i - \mathbf{m}_i)^T \right), \quad (3)$$

$$\mathbf{S}_b = (\mathbf{m}_1 - \mathbf{m}_2) (\mathbf{m}_1 - \mathbf{m}_2)^T.$$

The Fisher criterion is that the samples of the same class are aggregated in the Fisher vector space, while the samples of different class are separated as much as possible. So the objective function is defined as follows:

$$J_F(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}. \quad (4)$$

By Lagrange Multiplier, the Fisher vector to this maximization problem is as follows:

$$\mathbf{w} = \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2). \quad (5)$$

Then each skull feature vector  $\mathbf{a}$  can be projected to the Fisher vector, in which two classes of samples can be well separated:

$$t = \mathbf{W}^T \mathbf{a}. \quad (6)$$

Finally, a threshold is selected by using some prior knowledge, for example,

$$t_0 = \mathbf{W}^T \frac{(\mathbf{m}_1 + \mathbf{m}_2)}{2}. \quad (7)$$

Given an unknown skull, the procedure of the sex determination is as follows.

*Step 1.* Perform the statistical shape model matching to determine the shape parameter as Figure 3 describes.

*Step 2.* Project the shape parameter to the Fisher vector, and get the projection value  $t$ .

*Step 3.* If the value is larger than the threshold, the skull gender is male. Otherwise, it is female.

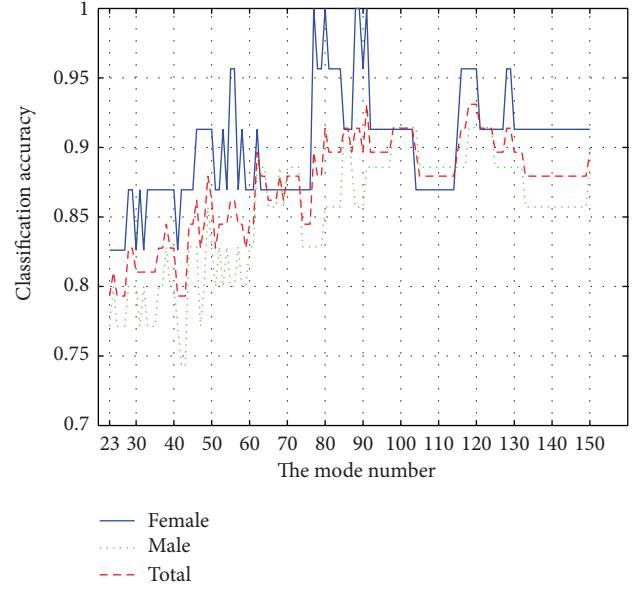


FIGURE 4: Variation of the correct classification rate of the test samples with the mode number.

### 3. Results

The used data are the 208 whole skull CT scans described in Section 2. We chose about three-fourths of the 208 skulls according to the age distribution as training samples, including 92 males and 58 females, and other skulls as test samples. In order to determine an appropriate mode number ( $M$  in (1)) to build the statistical shape model, we use the trial and error technique and analyze the variation of the classification accuracy of the test samples with varying the mode number by experiments. It is well known that, in PCA, generally only some eigenvectors corresponding to large eigenvalues of the covariance matrix represent modes of variation of the data, while others represent noises. We vary the mode number from 23 to 150, corresponding to variance contribution rate from 96% to 100%, and the correct classification rates for test samples are shown in Figure 4, while the rate of training samples is 100% when the mode number is larger than 80. From Figure 4 we can see that the correct classification rate of females is always higher than that of males, both of the correct rates for females and males are above 86% when the mode number is greater than 80, and the total correct rate for all test samples reaches maximum with the mode number being 91, 118, 119, and 120. Since the correct rate is stable with the mode number varying from 118 to 120, we use 119 modes to build the statistical shape model.

Figure 5(a) shows the classification of the training samples with the mode number being 119, and the classification accuracy of the training samples is 100%. Figure 5(b) shows the classification of the test samples. One female and 3 males are misclassified, and the correct rate is 95.7% and 91.4% for females and males, respectively. We also classify the whole 208 samples using leave-one-out strategy. That is, repeatedly 207 samples are chosen as training samples, and the residual



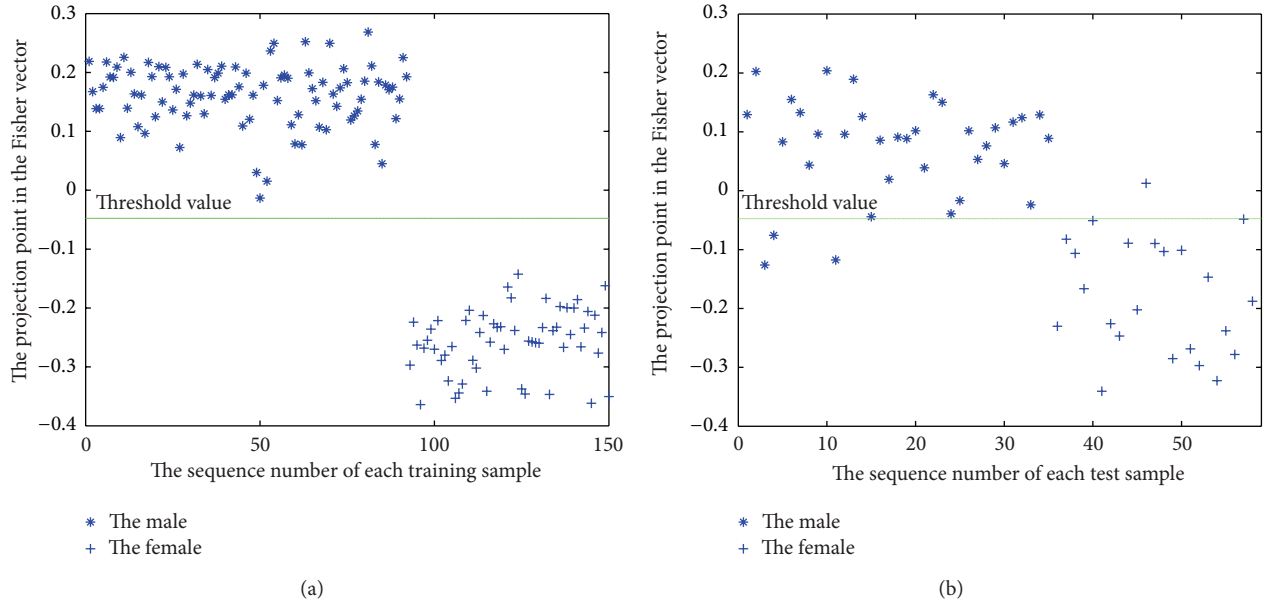


FIGURE 5: Classification of training samples (a) and test samples (b).

one is used as test sample. The mode number is still 119, which maybe is not the optimal selection. The correct classification rate is 91.3% and 90.1% for males and females, respectively. Compared with the accuracy of 87% by Shui [21], which is a metric analysis method and uses a subset of our 208 skulls to establish the discriminant functions, this method highly improves the accuracy.

#### 4. Discussion

Traditional morphological methods depend heavily on physical anthropologists' subjective understanding of population differences in sexual dimorphism. Different observers usually have a significant difference when performing the visual assessment of the morphological traits, especially for those inexperienced observers. On the other hand, discriminant analysis for skeletal measurements depends less on the examiner's professional qualification and experience, but it requires a high measurement precision, which is not easy to realize. Majority of scientists contend that the combination of metrical and morphological methods is the best way [28]. In this work, we use 3D dense point cloud data to build the statistical shape model, and in a sense, the model parameters metrically describe the skull morphology. So the method combines the advantages of metrical and morphological methods. Different from previous methods, what this method measures is not distance- or volume-related variables, but global shape variations.

Previous studies [18] show that the performance of sex determination outside the reference population group for which the discriminant function has been developed is poor. We do not know whether this method also has this problem due to lack of skull samples of another population, but there indeed exists the problem of selection of training samples. The more complete the training samples are, the better

the performance is. That is, the training samples should depict shape variations as much as possible. We think that only if other populations have similar shape variations, the discriminant model we build is applicable. That just is the reason that the correct classification rate of females is uniformly higher than that of males, as shown in Figure 4. When the statistical shape model and the discriminant function have been established, sex determination of an unknown skull becomes easy to perform, and needs no professional qualification and tedious manual measurement. Moreover, although we use CT scans to construct 3D point cloud model of the skull in this work, the statistical shape model we build also can deal with 3D models constructed in any way such as laser scan 3D camera.

#### Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (no. 61272363) and Program for New Century Excellent Talents in University (NCET-13-0051).

#### References

- [1] W. M. Bass, *Human Osteology: A Laboratory and Field Manual*, Missouri Archaeological Society Columbia, 5th edition, 2005.
- [2] D. L. France, "Observation and metric analysis of sex in the skeleton," in *Forensic Osteology: Advances in the Identification of Human Remains*, K. J. Reichs, Ed., pp. 163–186, Charles C. Thomas, Springfield, Ill, USA, 1998.
- [3] R. B. Pickering and D. C. Bachman, *The Use of Forensic Anthropology*, CRC Press, Boca Raton, Fla, USA, 1997.
- [4] D. H. Ubelaker and C. G. Volk, "A test of the Phenice method for the estimation of sex," *Journal of Forensic Sciences*, vol. 47, no. 1, pp. 19–24, 2002.

- [5] T. L. Rogers, "Determining the sex of human remains through cranial morphology," *Journal of Forensic Sciences*, vol. 50, no. 3, pp. 493–500, 2005.
- [6] P. L. Walker, "Sexing skulls using discriminant function analysis of visually assessed traits," *American Journal of Physical Anthropology*, vol. 136, no. 1, pp. 39–50, 2008.
- [7] Regina vJ-LJ, S.C.R. File No. 26830, 2000.
- [8] United States Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals Inc.* 509 U.S. 579, 1993.
- [9] B. A. Williams and T. L. Rogers, "Evaluating the accuracy and precision of cranial morphological traits for sex determination," *Journal of Forensic Sciences*, vol. 51, no. 4, pp. 729–735, 2006.
- [10] W. M. Krogman and M. Y. Iscan, *The Human Skeleton in Forensic Medicine*, Charles C. Thomas, Springfield, Ill, USA, 1986.
- [11] F. Ramsthaler, M. Kettner, A. Gehl, and M. A. Verhoff, "Digital forensic osteology: morphological sexing of skeletal remains using volume-rendered cranial CT scans," *Forensic Science International*, vol. 195, no. 1–3, pp. 148–152, 2010.
- [12] T.-H. Hsiao, H.-P. Chang, and K.-M. Liu, "Sex determination by discriminant function analysis of lateral radiographic cephalometry," *Journal of Forensic Sciences*, vol. 41, no. 5, pp. 792–795, 1996.
- [13] T.-H. Hsiao, S.-M. Tsai, S.-T. Chou et al., "Sex determination using discriminant function analysis in children and adolescents: a lateral cephalometric study," *International Journal of Legal Medicine*, vol. 124, no. 2, pp. 155–160, 2010.
- [14] M. Inoue, T. Inoue, Y. Fushimi, and K. Okada, "Sex determination by discriminant function analysis of lateral cranial form," *Forensic Science International*, vol. 57, no. 2, pp. 109–117, 1992.
- [15] S. A. Veyre-Goulet, C. Mercier, O. Robin, and C. Guérin, "Recent human sexual dimorphism study using cephalometric plots on lateral telerradiography and discriminant function analysis," *Journal of Forensic Sciences*, vol. 53, no. 4, pp. 786–789, 2008.
- [16] M. K. Spradley and R. L. Jantz, "Sex estimation in forensic anthropology: skull versus postcranial elements," *Journal of Forensic Sciences*, vol. 56, no. 2, pp. 289–296, 2011.
- [17] R. L. Jantz and S. D. Ousley, *FORDISC 3: Computerized Forensic Discriminate Functions*, Version 3.0, University of Tennessee, Knoxville, Tenn, USA, 2005.
- [18] P. Guyomarc'h and J. Bruzek, "Accuracy and reliability in sex determination from skulls: a comparison of Fordisc 3.0 and the discriminant function analysis," *Forensic Science International*, vol. 208, no. 1–3, pp. 180.e1–180.e6, 2011.
- [19] F. Ramsthaler, K. Kreutz, and M. A. Verhoff, "Accuracy of metric sex analysis of skeletal remains using Fordisc based on a recent skull collection," *International Journal of Legal Medicine*, vol. 121, no. 6, pp. 477–482, 2007.
- [20] C. B. Li, "Study on sex difference of adult human skull in the northest china by stepwise discriminant analysis," *Journal of China Medical University*, vol. 1, no. 21, pp. 28–31, 1992.
- [21] W. Y. Shui, *Research and application of craniofacial morphology [dissertation]*, Beijing Normal University, 2011.
- [22] F. Q. Duan, S. Yang, D. H. Huang et al., "Craniofacial reconstruction based on multi-linear subspace analysis," *Multimedia Tools and Applications*, 2013.
- [23] X.-H. Liu, Y. Sun, and Y. Wu, "Reduction of hyperspectral dimensions and construction of discriminating models for identifying wetland plant species," *Spectroscopy and Spectral Analysis*, vol. 32, no. 2, pp. 459–464, 2012 (Chinese).
- [24] Q. Q. Deng, M. Q. Zhou, W. Y. Shui, Z. Wu, Y. Ji, and R. Bai, "A novel skull registration based on global and local deformations for craniofacial reconstruction," *Forensic Science International*, vol. 208, no. 1–3, pp. 95–102, 2011.
- [25] Y. Hu, F. Duan, B. Yin et al., "A hierarchical dense deformable model for 3D face reconstruction from skull," *Multimedia Tools and Applications*, vol. 64, pp. 345–364, 2013.
- [26] Y. Hu, F. Duan, M. Zhou et al., "Craniofacial reconstruction based on a hierarchical dense deformable model," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, article 217, 14 pages, 2012.
- [27] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
- [28] B. Sejrsen, N. Lynnerup, and M. Hejmadi, "An historical skull collection and its use in forensic odontology and anthropology," *Journal of Forensic Odonto-Stomatology*, vol. 23, no. 2, pp. 40–44, 2005.